

RECEIVED

AUG 03 2001

1631

TECH CENTER 1600,2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001  
TIME: 10:00:10Input Set : A:\873726u1.sql  
Output Set: N:\CRF3\07112001\H939905B.raw5 110> APPLICANT: Glaxo, Mark  
9 120> TITLE OF INVENTION: Soybean Seed Coat Peroxidase Structural Gene And Regulatory

## Region

13 130> FILE REFERENCE: 75-105  
17 140> CURRENT APPLICATION NUMBER: US 08/939,905B  
C--> 19 <141> CURRENT FILING DATE: 1997-09-29

23 150&gt; PRIOR APPLICATION NUMBER: US 08/778,414

25 151&gt; PRIOR FILING DATE: 1996-09-30

29 160&gt; NUMBER OF SEQ ID NOS: 19

33 170&gt; SOFTWARE: PatentIn version 3.0

37 180&gt; SEQ ID NO: 1

39 211&gt; LENGTH: 1244

41 212&gt; TYPE: DNA

43 213&gt; ORGANISM: Glycine max

47 220&gt; FEATURE:

49 221&gt; NAME/KEY: CDS

51 222&gt; LOCATION: (1)..(1056)

55 220&gt; FEATURE:

57 221&gt; NAME/KEY: sig\_peptide

59 222&gt; LOCATION: (1)..(78)

63 400&gt; SEQUENCE: 1

64	atg ggt tcc atg cct cta tta gta gag gaa ttg ttg tgt gca ttt gct	48
65	Met Gly Ser Met Ala Leu Leu Val Val Ala Leu Leu Cys Ala Phe Ala	
66	1 5 10 15	
68	atg cat gca ggt ttt tca gtc tct tat gct cag ctt act cct acg ttc	96
69	Met His Ala Gly Phe Ser Val Ser Tyr Ala Gln Leu Thr Pro Thr Phe	
70	20 25 30	
71	tac aga gaa aca tct cca aat ctg ttc cct att gtg ttt gga gta atc	144
72	Tyr Arg Glu Thr Cys Pro Asn Leu Phe Pro Ile Val Phe Gly Val Ile	
73	35 40 45	
74	ttc gat gct tct ttc acc gat ccc cga atc ggg gcc agt ctc atg agg	192
75	Phe Asp Ala Ser Phe Thr Asp Pro Arg Ile Gly Ala Ser Leu Met Arg	
76	50 55 60	
78	cct cat ttt cat gat tgc ttt gtt taa ggt tgt gat gga tca gtt tgg	240
79	Leu His Phe His Asp Cys Phe Val Phe Gly Cys Asp Gly Ser Val Leu	
80	65 70 75 80	
81	ctg aac aac act gat aca ata gaa agc gag caa gat gca ctt cca aat	288
82	Leu Asn Asn Thr Asp Thr Ile Glu Ser Glu Gln Asp Ala Leu Pro Asn	
83	85 90 95	
84	atc aac tca ata aga gga ttg gac att gtc aat gac atc aag aca gag	336
85	Ile Asn Ser Ile Arg Gly Leu Asp Val Val Asn Asp Ile Lys Thr Ala	
86	100 105 110	
87	gtg gaa aat agt tgt cca gac aca gtt tct tgt gct gat att ctt gct	384
88	Val Glu Asn Ser Cys Pro Asp Thr Val Ser Cys Ala Asp Ile Leu Ala	
89	115 120 125	
90	att gca gct gaa ata gct tct gtt ctg gga gga ggt cca gga tgg cca	432
91	Ile Ala Ala Glu Ile Ala Ser Val Leu Gly Gly Gly Pro Gly Trp Pro	
92	130 135 140	

ENTERED

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001  
 TIME: 10:0 :10

Input Set : A:\873726u1.sq1  
 Output Set: N:\CRF3\07112001\H939905B.raw

```

100 gtt cca tta gga aga agg gac ays tta ara gca aac cga acc ctt gca 480
101 Val Pro Leu Gly Arg Arg Asp Ser Leu Thr Ala Asn Arg Thr Leu Ala
102 145 150 155 160
104 aat caa aac ctt cca gaa cct tta tta aac cta act caa ctt aaa gct 528
105 Asn Gln Asn Leu Pro Ala Pro Phe Phe Asn Leu Thr Gln Leu Lys Ala
106 165 170 175
108 tcc att gct gtt caa ggt ctc aac acc cta gat tta gtt aca ctc tca 576
109 Ser Phe Ala Val Gln Gly Leu Asn Thr Leu Asp Leu Val Thr Leu Ser
110 180 185 190
112 ggt ggt cat acc ttt gga aga gct cag tgc agt aca ttc ata aac cga 624
113 Gly Gly His Thr Phe Gly Arg Ala Arg Cys Ser Thr Phe Ile Asn Arg
114 195 200 205
116 tta taa aac tta agt aac act gga aac cct gat cca act ctg aac aca 672
117 Leu Tyr Asn Phe Ser Asn Thr Gly Asn Pro Asp Pro Thr Leu Asn Thr
118 210 215 220
120 aca tac tta gaa gta tgg cgt gaa aga tta cca cag aat gaa act ggt 720
121 Thr Tyr Leu Glu Val Leu Arg Ala Arg Cys Pro Gln Asn Ala Thr Gly
122 225 230 235
124 gat aac ctc acc aat ttg gac ctg agc aca cct gat caa ttt gac aac 768
125 Asp Asn Leu Thr Asn Leu Asp Leu Ser Thr Pro Asp Gln Phe Asp Asn
126 240 245 250 255
128 aga tac tac tta aat ctt ctg tag ctc aat gac tta ctt cag agt gac 816
129 Arg Tyr Tyr Ser Asn Leu Leu Gln Leu Asn Gly Leu Leu Gln Ser Asp
130 260 265 270
132 caa gaa ctt tta tta act cct ggt cct gat acc att ccc att gtc aat 864
133 Gln Glu Leu Phe Ser Thr Pro Gly Ala Asp Thr Ile Pro Ile Val Asn
134 275 280 285
136 agc ttc agc agt aac cag aat act ttc ttt tta aac ttt aga gtt tra 912
137 Ser Phe Ser Ser Asn Gln Asn Thr Phe Phe Ser Asn Phe Arg Val Ser
138 290 295 300
140 atg ata aaa atg ggt aat att gga gtg ctg act ggg gat gaa gga gaa 960
141 Met Ile Lys Met Gly Asn Ile Gly Val Leu Thr Gly Asp Glu Gly Glu
142 305 310 315 320
144 att ggc tgg caa tgg aat ttt gtg aat gga gac tgg ttt gga tta gct 1008
145 Ile Arg Leu Gln Cys Asn Phe Val Asn Gly Asp Ser Phe Gly Leu Ala
146 325 330 335
148 agt gtg ggc tcc aaa gat gct aaa caa aag ctt gtt gct caa tct aaa 1056
149 Ser Val Ala Ser Lys Asp Ala Lys Gln Lys Leu Val Ala Gln Ser Lys
150 340 345 350
152 taaaccaata attaatgggg atgtgatgc tagctagcat gttaaaggcaa attaggttgt 1116
154 aaacctcttt gctagctata ttgaaataaa ccaaaggagt agtgtgcatg tcaattcgat 1176
156 tttgcatgt accctctgga atattatgta ataattattt gaattcttt aaggtactta 1236
158 attaatca 1244
161 -110- SEQ ID NO: 1
163 -111- LENGTH: 4770
165 -112- TYPE: DNA
167 -113- ORGANISM: Glycine max
171 -120- FEATURE:
173 -121- NAME/KEY: promoter

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001

TIME: 10:00:10

Input Set : A:\873726u1.sql

Output Set: N:\CRF3\07112001\H939905B.raw

```

175 - 222 - LOCATION: (1)..(1532)
176 - 222 - FEATURE:
177 - 221 - NAME/KEY: sig_peptide
178 - 222 - LOCATION: (1733)..(1610)
179 - 222 - FEATURE:
180 - 221 - NAME/KEY: exon
181 - 222 - LOCATION: (1733)..(1751)
182 - 222 - FEATURE:
183 - 221 - NAME/KEY: exon
184 - 222 - LOCATION: (2333)..(2574)
185 - 222 - FEATURE:
186 - 221 - NAME/KEY: exon
187 - 222 - LOCATION: (2605)..(2769)
188 - 222 - FEATURE:
189 - 221 - NAME/KEY: exon
190 - 222 - LOCATION: (4033)..(4512)
191 - 222 - FEATURE:
192 - 221 - NAME/KEY: Intron
193 - 222 - LOCATION: (1152)..(2383)
194 - 222 - FEATURE:
195 - 221 - NAME/KEY: Intron
196 - 222 - LOCATION: (2375)..(3604)
197 - 222 - FEATURE:
198 - 221 - NAME/KEY: Intron
199 - 222 - LOCATION: (3670)..(4012)
200 - 222 - FEATURE:
201 - 221 - NAME/KEY: CDS
202 - 222 - LOCATION: (1513)..(1751)
203 - 222 - FEATURE:
204 - 221 - NAME/KEY: CDS
205 - 222 - LOCATION: (1383)..(2574)
206 - 222 - FEATURE:
207 - 221 - NAME/KEY: CDS
208 - 222 - LOCATION: (2605)..(2769)
209 - 222 - FEATURE:
210 - 221 - NAME/KEY: CDS
211 - 222 - LOCATION: (4033)..(4512)
212 - 400 - SEQUENCE:
213 tagataaaaaa atgggparat aatttttctc agatgttggt tatactgttt ttttaatcag 60
214 aattaaaaat cttctttaat tatcgacata atttttttgg gtgaatatta tcgacataat 120
215 tatttaaaa aaatttttat tjtacataga agtgataact caattttaat attggagaac 180
216 agtaagaaaa attaaaaaaa cgttatttag aagaaaaaaa tatatggaaa aggttagcta 240
217 catatattat aaaaatagt tgtttaatt ggctatataa acctattgt actctttgta 300
218 atttcacatt attcatttaa ataatcttct actttttaag ttctatattt tctctcaait 360
219 ttctctgata acccaaaaaa tttaacatgg tatatcagcg ataccaccca ctttgaaaagc 420
220 catatatggt tagtatggc agcaaaatt tgccctgggt caagcaaaagc aagtgtttat 480
221 atagatgtga cttttgtga ggaatcagc ccaatggtae tgattgtgaa actgagaaaa 540
222 aaatttggg gaatttgaat tatgatcatt aaatactcct ctctgacta ccttcgtccc 600
223 taaatttjt accaatttta ttcccaaaaa atttgattac aatgcactaa ttaatgaatg 660

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001

TIME: 19:00:10

Input Set : A:\873726u1.sql

Output Set: N:\CRF3\07112001\H939905B.raw

```

293 ttttttacat ttttttacat ttttttacat ttttttacat ttttttacat ttttttacat 720
300 ttttttttgt ttttttttgt ttttttttgt ttttttttgt ttttttttgt ttttttttgt 731
302 aaatgcatga aaatgcatga aaatgcatga aaatgcatga aaatgcatga aaatgcatga 840
304 tgattatttt ttgcaaatga atgtttatgt aacattttaa ttttttttgt ttttttttgt 901
306 ttatggtgtc atgtttttaa aacattttaa aacattttaa aacattttaa aacattttaa 91
308 atttttaaact ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1019
310 ttttttttgt ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1030
312 aaccatatta aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1140
314 ggttttgtgt ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1200
316 ttgttttgtt ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1210
318 aatgatttgt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1220
320 ttgttttgtt ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1230
322 ttgttttgtt ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1440
324 aatgatttgt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1450
326 aatgatttgt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1553
328 aatgatttgt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1561
330 gta gta gta gta gta gta gta gta gta gta gta gta gta gta gta gta 1571
331 Val Val Ala Leu Leu Cys Ala Phe Ala Met His Ala Gly Phe Ser Val 1581
332 10 15 20 1591
334 ttt tat gct gaa gta gta gta gta gta gta gta gta gta gta gta gta 1597
335 Ser Tyr Ala Gln Leu Thr Pro Thr Phe Tyr Arg Glu Thr Cys Pro Asn 1601
336 25 30 35 1607
338 atg ttc cct att gta gta gta gta gta gta gta gta gta gta gta gta 1611
339 Leu Phe Pro Ile Val Phe Gly Val Ile Phe Asp Ala Ser Phe Thr Asp 1621
340 40 45 50 55 1631
342 ccc cga atc gga gga gga gga gga gga gga gga gga gga gga gga gga 1641
343 Pro Arg Ile Gly Ala Ser Leu Met Arg Leu His Phe His Asp Cys Phe 1651
344 60 65 70 1661
346 gtt caa gta gta gta gta gta gta gta gta gta gta gta gta gta gta 1671
347 Val Gln 1681
349 ttttttgttt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1691
351 ttttttgttt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1701
353 aaggtatttt ttgcaaatga aacattttaa aacattttaa aacattttaa aacattttaa 1711
355 ttgaatgata ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1721
357 gaagaaaaaa gatgttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1731
359 taattaatac ttttttttgt aacattttaa aacattttaa aacattttaa aacattttaa 1741
361 gacaagtatt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1751
363 ttgtatagat aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1761
365 tacattgatt aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1771
367 tgatctgaac aacattttaa aacattttaa aacattttaa aacattttaa aacattttaa 1781
369 75 1791
371 gtt ttg ctg aac aac act gta aca ata gaa aac gag caa gat gaa ctt 1801
373 Val Leu Leu Asn Asn Thr Asp Thr Ile Glu Ser Glu Gln Asp Ala Leu 1811
374 80 85 90 1821
376 cca aat atc aac tca ata aga gga ttg gac gtt gtc aat gac atc aag 1831
377 Pro Asn Ile Asn Ser Ile Arg Gly Leu Asp Val Val Asn Asp Ile Lys 1841
378 95 100 105 110 1851

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001

TIME: 19:00:10

Input Set : A:\873726u1.sql

Output Set: N:\CRF3\07112001\H939905B.raw

```

380 aca ggg gtg gaa aat agt tgt cca gac aca gtt tct tgt gct gat att 2541
381 Thr Ala Val Glu Asn Ser Cys Pro Asp Thr Val Ser Cys Ala Asp Ile
382 115 120 125
384 ttt gct att gca ggt gaa ata gct tct gtt ctg gtaattasta actcctaatt 2594
385 Leu Ala Ile Ala Ala Glu Ile Ala Ser Val Leu
386 130 135
388 gattcccaac cattaaaaag ttgcattgatt ggattcaaaa ttttatgata ttgggggttct 2694
389 gatataaatt tgaattaaa ttgcattaaa aaaaattat atatacttt aataaaaaaa 2714
390 atttatctaa ttaatttat tattaaaaat atttttaaaa ttcaattcta actctttttt 2734
391 atcgggagca tgaagctgg caccacacgt atatcgttg aagatgcat aaaaaccaat 2834
392 aattaatgga tggatcact caaaacattt aattcaaaa actottaatt gtgattagta 2894
393 atcatgttcg ggcagtttac gtttgtgata attaatltga cttaataaga taaaaaaaca 2994
394 atggagcga agcgggttgg tatagatct actgggctgt agaatat gtg atttttcaag 3014
395 tttaaataaa agttactat tatattatat ttagtctttt tttttctaa acccatttaa 3074
396 agtgatttat tgaattgaa acatgttct acacacaggg ttgaaaactc ctggcaacta 3134
397 acatctccaa aatttcacaa ttatatatg aagataattt atctatgctg ttaactcta 3194
398 ttatatatat gtatcatcgc agtattcaga attataatag tcaaatatag agtatatcg 3214
399 ggtaaatgta gttgcattgg agactgttt cgtctaaaa gottattcta tatagctttt 3314
400 tttattggaa aataacgat aactaaaaac gaaagggtat catatgttt gacttttatg 3374
401 tttagagagag acatctaat ttggtcatat gttaaataat taattacaat gcatacaca 3434
402 atatttatgc catctcaaaa aaatgataaa atatcatagg tatactcaac tatatgat 3494
403 ccccataaca aaatttgtac tttttctcag gcaatgaact taacattct gtttgctaaa 3554
404 aacaaacatc aatttaaaat ggttcaaat atttatgtaa taatttcag gga gga 3610
405 3659
406 ggt cca gga tgg cca gtt cca tta gga aca agg gac agt tta aca gca
407 Gly Pro Gly Trp Pro Val Pro Leu Gly Arg Arg Asp Ser Leu Thr Ala
408 140 145 150 155
409 aac cga acc att gca aat caa aac ctt cca gca cct ttc ttc aac ctg
410 Asn Arg Thr Leu Ala Asn Gln Asn Leu Pro Ala Pro Phe Phe Asn Leu
411 160 165 170
412 act caa ctt aaa gct tca ttt gct gtt caa ggt ctg aac acc att gat
413 Thr Gln Leu Lys Ala Ser Phe Ala Val Gln Gly Leu Asn Thr Leu Asp
414 175 180 185
415 tta gtt aca ctg tca ggtatacata atcaattttt tatttgcctat tagctagcaa 2809
416 Leu Val Thr Leu Ser
417 190
418 taaaaagtcg ctgatacaga catattttaga taaattaat tctccataaa catttataat 2869
419 aaaattatca atttatgtac ttaaaaaatta tggattgaag ctctttcat caaactttta 2929
420 cttaagttaa ggtgcattata atataaaata aacratctct tgtttttat aaaaagattg 2989
421 aagataagtt aagctctact tataaatcat taatalangt ata ggt ggt cat acg 3044
422 315
423 gga aga cct cgg tgc agt aca ttc ata aac cga tta taa aac ttc 3092
424 Phe Gly Arg Ala Arg Cys Ser Thr Phe Ile Asn Arg Leu Tyr Asn Phe
425 200 205 210
426 agc aac act gga aac cct gat cca act ctg aac aca aca tac tta gaa 4140
427 Ser Asn Thr Gly Asn Pro Asp Pro Thr Leu Asn Thr Thr Tyr Leu Glu
428 215 220 225
429 gta ttg cgt gca aga tgc ccc cag aat gca act ggg gat aac ctg acc 4158

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001

TIME: 10:00:11

Input Set : A:\873726u1.sql

Output Set: N:\CRF3\07112001\H939905B.raw

L:19 M:271 C: Current Filing Date differs, Replaced Current Filing Date